

## Index

- #LancsBox, 25, 224, 226, 227
- acceptability, 169
- admitted hypothesis, 60
- AmE06  
   replication of, 226
- AmE06 Corpus, 222
- AmE06' Corpus, 222
- AmE06" Corpus, 222
- American English, 125  
   corpus-based study of, 128  
   modal verbs in, 218  
   spoken, 181
- American v. British English  
   decline of modals over time, 230  
   frequencies of modals over time, 233–234  
   high and low frequency modals over time, 231–232  
   modal rank correlation, 238  
   over time, 214  
   the case of *could* and *might*, 235  
   the decline of core modals, 244–245  
   the stability of high and low frequency modals over time, 232–233  
   volatility in modal frequency over time, 236–237
- ampliative reasoning, 20, 27, 138
- annotation  
   and ontology, 16, 182, 191  
   and underdetermination, 182, 183
- anticipation, 98, 120, 136
- anti-naturalism, 113, 124
- a-priori assumption, 15
- Aristotelian  
   essences, 122
- Assamese, 171
- assumption of purposefulness, 159, 161, 162
- axiom  
   and corpora, 74  
   and descriptive and explanatory adequacy, 76–77  
   and extra-logical constants, 74, 75  
   and interdependency, 92, 177  
   and precision, 103  
   and range, 96  
   and stray statements, 91  
   by convention, 76, 85, 96  
   in chemistry, 166  
   of a theory, 74  
   of collocation, 74, 77, 80  
   of narrative, 91, 92  
   of quotation, 91, 92  
   of transcription, 176–178
- axiom by convention, 78, 224
- AC1, 225
- AC2, 225–226
- AC3, 226
- AC4, 226
- AC5, 227
- AC6, 233, 234–235
- AC7, 237
- AC8, 239, 241
- axiomatic functionalism, 160–161  
   origins, 160
- balance  
   in corpus data, 27, 250  
   of propensity, 246
- basic statement, 72  
   incremental refinement, 90  
   revision of, 90  
   role in theory, 80
- Bayesian reasoning, 138

- BE06
  - replication of, 226
- BE06 Corpus, 222
- BE06' Corpus, 222
- BE06" Corpus, 222
- BNC 1994
  - demographic component, 189, 195
- breach
  - of principle, 84, 162
- British English, 125
  - corpus-based study of, 128
  - modal verbs in, 218
- Brown corpus, 102
- Brown family of corpora
  - and falsification, 64, 102
  - and propensity, 147
  - and the exploration of genre, time and variety, 204
  - development of sampling structure, 225
  - extension, 199
  - part-of-speech tagging of, 191, 199, 224
  - reconfiguration, 214
  - sampling frame, 64, 65, 249
- Cartesian, 11
- causation, 138–139, 246, 248–249
  - and language variety, 213
  - distal, 246
  - proximal, 246
- certainty
  - and ampliative reasoning, 20
  - and induction, 39
  - and justificationism, 62–63
  - and Newtonian physics, 39
  - and scientific élites, 50
  - and the management of uncertainty, 111
  - and the rationality principle, 118
- Cartesian dream of, 11
- dangers of, 152
- unattainable, 178
- change over time
  - and corpus construction, 147
  - and critical realism, 189
  - and study design in the social sciences, 202–203
  - diachronic falsification, 137–138
  - Chinese, 58
- Chomsky, Noam
  - and naturalism, 130
  - and scientia rationalis, 23
  - career, 7
  - intuition and dogmatism, 54, 58, 61
  - intuition and reality, 50–52
  - Key Thinker entry, 7
  - universalism v. axiomatic functionalism, 161
  - views on corpus linguistics, 6–7
- cohort studies
  - and corpus linguistics, 202
  - and matched samples, 226
- common sense, 29
  - and observation, 12–13
  - and rationality, 10
  - in science, 11
  - view of reality, 10
  - view of science, 9–10
- consistency, 82, 83
- construction grammar, 73, 74, 182
- convention, 64, 65, 66, 67, 71, 76, 87, 90, 97, 106, 109, 110, 118, 123, 149, 183, 189, 191, 192, 197, 237
  - and a-priorism, 107
  - and bias, 150
  - and dogmatism, 68, 90–91
  - and extra-logical constants, 75
  - and statement equations, 75
  - propensity and statistics, 240–244
- conventionalism. *See* anti-naturalism
- conviction, 49, 59
  - and inter-subjective testing, 60
  - and objection, 71
  - and systems of belief, 86
- corpus internal evidence, 168, 172
  - and acceptability, 176
- corpus-based, 31, 85, 92, 185
- corpus-driven, 31, 93
  - problems with, 41–42
- corroboration, 35, 65, 79, 83, 94, 97, 205
  - and convention, 90
  - and conviction, 59
  - and fallibilism, 152
  - and probability, 100, 101–105

## Index

309

- and provisional truth, 75
- and statistics, 250
- degrees of, 101
- CQPweb, 186, 192
- critical rationalism, 131, 132, 166
  - and corpus data, 173
  - and probability, 138
- critical realism, 10, 47, 132
  - and change over time, 189
  - and corpus linguistics, 119
  - and critical rationalism, 133–134
  - and modelling, 251
  - and naive realism, 132–133
- cross-sectional studies, 202, 205
  - design, 204
- critical rationalism
  - and critical realism, 133–134
- culturomics, 103, 113
- dark matter, 150–151
- deduction, 17, 36, 106
  - and quasi-induction, 106–107
- deeming, 63, 89, 90, 94, 98, 105
  - and convention, 90–91
- degree of falsifiability, 95
- deterministic, 133, 140
- doctrine of revisability, 89
- dogmatism, 58, 60, 61, 67, 90, 98, 125, 131, 167
  - and justificationism, 63
- Dušková, Libuše
  - Key Thinker entry, 158
- effect size, 53, 237, 243
- eliminative inference, 65
- empirical content, 95, 96, 100, 128, 137, 177, 210, 221
- empirical range, 137, 138
  - and surprise, 222
- empiricism, 22–25
- epistemology, 15, 16, 152
  - and presupposition, 16
  - and statistical tests, 241
- distinctions, 23
- functionalist, 157
- normative, 89, 171, 174, 207
- Popper's and Wittgenstein's
  - contrasted, 253
  - position of this book, 131, 254–255
  - social-scientific, 217
- essentialism, 122
- existential statement, 72, 207
  - inverse, 72, 206
- experimental design, 198, 201, 203–204, 215
- experimental falsifiers, 99, 128
- explanatory power, 77, 126, 129
- external explanation, 155
- extra-scientific, 116
  - link to the scientific, 116
  - results, 115
- fallibilism, 109, 112
  - and corroboration, 152
  - and scepticism, 109
- falsifiability, 42–45, 79, 123
  - and methodology, 60
  - and simplified data, 127
  - degree of, 79, 96, 99, 102, 105
- falsifiable, 43, 44, 75
  - in correspondence with reality, 76
- falsification, 80
  - and ampliative reasoning, 100
  - and conditional falsification, 107–108
  - and conviction, 58
  - and corroboration, 45
  - and diachrony, 136
  - and dogmatism, 61, 71
  - and empirical content, 96
  - and intuition, 49
  - and persistence, 79
  - and prototypicality, 158
  - and range, 97
  - and rationality, 130
  - and recursion, 109
  - and rejection, 108, 129, 167
  - and testability, 93
  - and the ordering of statements, 56
  - avoiding, 57
  - conditional, 108, 159, 206, 207, 245
  - in-situ, 163
  - in-situ, and deception, 166

## 310

falsification (cont.)  
   methodological, defined and discussed, 243  
 falsifiers  
   potential, 80  
 FLOB corpus, 128  
 flout, 162, 166, 168, 175  
 focus on the individual, 119–120, 121, 130,  
   135, 162, 169, 175  
   and probability, 141–142  
 formalism  
   versus functionalism, 154–155  
 frequency, 48, 73, 94, 169  
   and propensity, 141  
   word, 44, 66, 145, 180–182, 207  
 frequentist, 243  
 Frown corpus, 192  
 functionalism, 156  
   and performance data, 162  
   and the Prague School, 157  
   axiomatic. *See* axiomatic functionalism  
   Hallidayan, 159  
   v. formalism, 154–155  
 genuine test, 49  
 God's truth linguistics, 183  
   and annotation, 182  
   and the rationality principle, 118  
 God's truth linguistics, 118  
 gravity  
   and dark matter. *See* dark matter  
 Great Interlocutor, 43–44, 95  
 Hajičová, Eva  
   Key Thinker entry, 158  
 Halliday, Michael  
   Key Thinker entry, 158  
 hocus-pocus linguistics, 118  
 hypothesis  
   formation, 29, 75, 107, 247  
 idealism, 10  
 incremental refinement, 64, 97, 178  
 incrementalism, 90  
 indeterminism, 122, 140, 246  
   objective, 132  
 individual names, 69

## Index

induction, 17–18, 36, 37, 67, 112  
   and data scarcity, 18  
   and quasi-induction, 106  
   and recursion, 40  
   problem of, 37–41, 42, 99, 184  
 infinite regress, 88, 89, 90  
 inflectional paradigms, 52–54, 61  
 innate predispositions. *See* anticipation  
 institution, 119, 120  
   and interpretation, 121  
   and science, 123  
   interaction with traditions, 121  
   of language, 120, 130, 134, 135, 136  
   of language versus tradition, 121  
 instrumentation, 44, 90  
 intended audience, 165, 169, 170, 171  
 inter-subjective testing, 26, 59, 84, 86, 104,  
   130, 197  
   and conviction, 60  
 invariantist, 155  
 irrationalism, 49  
 justificationism, 63  
 Leech, Geoffrey  
   Key Thinker entry, 32  
 LOB corpus, 102  
 log-likelihood, 189, 192  
   and methodological falsification, 242  
 machine learning, 40  
 mathematics, 89, 116, 237, 242  
 metadata, 145  
   in the BNC 1994, 186  
   in the BNC 2014, 136  
 metaphysical, 44, 55, 56, 197, 218  
   addressing in corpus linguistics, 185  
   and anti-naturalism, 113  
   and replication, 221  
   and rules of acceptance, 59  
   and transition to the scientific, 44–45  
   Chomskyan use of intuition, 83–84  
   mutability of, 151  
   statement, 43  
 methodological  
   approaches, 14

## Index

311

- relativism, 10
- scepticism, 10
- methodology
  - and falsification, 60
  - and triangulation, 80
  - of science, 32
- modal verbs
  - declining, 11
  - hypothesis 1, conclusion, 231
  - hypothesis 2, conclusion, 231
  - hypothesis 3, conclusion, 235
  - hypothesis 4, conclusion, 235
  - hypothesis 5, conclusion, 238
  - hypothesis 6, conclusion, 239
  - hypothesis 7, conclusion, 239
  - hypothesis 8, conclusion, 239
  - study of, 218
- monism, 114
- Mulder, Jan
  - Key Thinker entry, 160
- natural science, 27, 36, 113, 245
  - and propensity, 149
  - and social science, 156
- naturalism, 136, 166
  - and Chomsky, 130
  - in the social sciences, 165
  - rejection of, 111–113
  - replacement, 245, 249
- new statistics, 243, 245
- no prior selection, 49, 83, 84
- nominalism
  - and linguistics, 122
- non-normative, 169
- no-prior selection, 84
- normative, 202
  - and replication, 245
  - and well-formedness, 179
  - conjecture, 169–170, 174, 175, 177
  - spelling, 125
  - view of data, 154
- numerical universality, 66
  - and corpora, 69
- observability, 88
- Ockham, William
  - Key Thinker entry, 23
- Oedipus effect, 125–126
- ontological knowledge, 16, 69, 70, 201, 210, 213
  - in linguistics, 70–72, 81
- ostensive definition, 76, 77
- particularist, 47, 50
- Popper, Karl
  - Key Thinker entry, 5
- positivism, 112
- potential falsifiers, 83, 85, 94, 96, 99, 137
  - and consistency, 83
  - and empirical content, 95
  - and ontological knowledge, 81
  - and time, 81
- poverty of stimulus, 61–62
- Prague Linguistic Circle
  - Key Thinker entry, 156
- precision, 103, 104
- premise, 17–18, 56
- presupposition. *See* a-priori assumption
- probability, 240, 242
  - and ampliative reasoning, 20
  - and corroboration, 100, 101–105
  - and critical realism, 216
  - and empirical content, 95
  - and propensity, 138–152, 241
  - and reality, 241
  - and significance statistics, 241
  - crystallised, 199
- propensity
  - and critical realism, 215
  - and probability, 138–152, 241
  - and sampling structure, 249
  - and volition, 247
  - convention and statistics, 240–244
  - of language, 184
  - variation of, 207, 210
- protocol sentence, 84
- provisional truth, 63, 68–72
- purposeful
  - nature of language, 157, 162, 179
- quasi-reality, 122, 129, 147, 149

## 312

range, 96–98  
 rationality principle, 126–128, 165, 172  
 realism, 9, 11, 12, 132  
 reality, 13, 15, 47, 87, 118  
   and intuition, 50–52  
   and model of reality, 69, 127, 129, 159  
   and probability, 241  
   and propensity, 147  
   and repeatability, 190  
   and triangulation, 134  
   approaches to, 9–11  
   correspondence to, 46–47, 50, 74  
   correspondence to and range, 96  
   correspondence to and truth, 12  
   correspondence to and realism, 12  
   critical approach to. *See* critical realism  
   degree of correspondence to, 96  
   individuals and, 119  
   linguistic, 12, 13, 16, 20, 55, 96, 110, 132, 161, 171, 184  
   linguistic, problematizing, 52  
   linguistic, quasi-contact with, 50  
   observed, 29  
   physical, 112  
   psychological, 56, 57, 77  
   quasi-contact with, 47, 48, 77  
   social, 15, 112, 131–132, 134  
 redundancy, 91  
 reification, 159  
   Chomskyan, 159  
 rejection, 159  
   in situ, 165  
 repeatability, 25–26, 69, 185–190, 195, 197  
   and broad replication, 217  
   and direct repetition, 217  
   and direct replication, 217  
   and repetition with variation, 217  
 repetition. *See* repeatability  
 replicability, 25–26, 200–201  
 replication  
   broad replication, 217, 218, 221, 222  
   direct replication, 217, 218  
   of the BE06 and AmE06 corpora, 224, 225  
 representativeness

## Index

in corpus data, 250, 251  
 research design, 201–206, 212  
 sampling, 201–203, 206  
   frame, 64, 207, 208, 212, 214, 226  
   oversampling, 250  
 sampling structure. *See* sampling frame  
 science  
   and paradigm, 20, 64, 150  
   Newtonian, 38  
   observational, 6, 66  
   philosophy of, 3, 33, 47, 99  
 scientia  
   rationalis, 22, 113  
   realis, 22, 24, 30  
 scientific  
   (to be), 6, 12, 14–15, 31–32, 35, 42, 43, 44, 45, 55, 60, 65, 68, 73, 82, 86, 109, 116, 117  
   and dogmatism, 60  
   transition from the metaphysical, 44–45  
   observations, 26  
   social, 124, 148, 149, 150, 197, 213, 250  
 scientific method, 8, 15, 22, 25, 27, 31, 34, 67, 92–93  
 searchlight theory, 93, 178, 185, 244  
 set of permitted statements, 80, 81, 96, 100  
 severity of a test, 101  
 significance testing, 240–242  
   and convention, 64  
   validity, 243  
 simplicity, 99–100  
 simplifying assumptions, 126, 167  
   in linguistics, 127  
 social context, 135  
   and language, 114, 121, 130  
   of language, 134  
 social science, 113, 137, 152  
   and causality, 248  
   and critical rationalism, 131  
   and natural science, 156  
   and naturalism, 165  
   and propensity, 152  
   and simplifying assumptions, 126  
   and the individual, 119

## Index

313

- epistemology, 217
- quantitative, 245, 246
- studies over time, 202–203
- statement equation, 74, 75
- statistical significance. *See* significance testing
- stray statement. *See* axiom, and stray statements
- Stubbs, Michael
  - Key Thinker entry, 31
- surprise
  - and corroboration, 102
  - and corpus data, 138
  - and empirical range, 222
  - and replication, 218, 221, 244
  - diachronic, 137
- test corpus, 200
- testability, 94, 95, 100
- topic modelling, 103–104
- total accountability, 49, 83, 84, 85
- tradition, 121, 135
  - and falsifiability, 123
  - and social context, 135
  - of language, 136–137
  - of language versus institution, 121
- training corpus, 200
- treatment design, 203–204
- underdetermination, 48, 182
- universal concepts, 69–70
- universal statement, 37, 38
- version management, 186, 194, 195
- Wittgenstein, Ludwig
  - Key Thinker entry, 253
- wordlists, 180–182